

# Quantum communication complexity of block-composed functions

Yaoyun Shi and Yufan Zhu

Department of Electrical and Computer Engineering

The University of Michigan

Ann Arbor, MI 48109-2122, USA

E-mail: {shiyy|yufanzhu}@eecs.umich.edu

April 14, 2008

## Abstract

A major open problem in communication complexity is whether or not quantum protocols can be exponentially more efficient than classical ones for computing a *total* Boolean function in the two-party interactive model. The answer appears to be “No”. In 2002, Razborov proved this conjecture for so far the most general class of functions  $F(x, y) = f_n(x_1 \cdot y_1, x_2 \cdot y_2, \dots, x_n \cdot y_n)$ , where  $f_n$  is a *symmetric* Boolean function on  $n$  Boolean inputs, and  $x_i, y_i$  are the  $i$ 'th bit of  $x$  and  $y$ , respectively. His elegant proof critically depends on the symmetry of  $f_n$ .

We develop a lower-bound method that does not require symmetry and prove the conjecture for a broader class of functions. Each of those functions  $F(x, y)$  is the “block-composition” of a “building block”  $g_k : \{0, 1\}^k \times \{0, 1\}^k \rightarrow \{0, 1\}$ , and an  $f_n : \{0, 1\}^n \rightarrow \{0, 1\}$ , such that  $F(x, y) = f_n(g_k(x_1, y_1), g_k(x_2, y_2), \dots, g_k(x_n, y_n))$ , where  $x_i$  and  $y_i$  are the  $i$ 'th  $k$ -bit block of  $x, y \in \{0, 1\}^{nk}$ , respectively.

We show that as long as  $g_k$  itself is “hard” enough, its block-composition with an *arbitrary*  $f_n$  has polynomially related quantum and classical communication complexities. Our approach gives an alternative proof for Razborov’s result (albeit with a slightly weaker parameter), and establishes new quantum lower bounds. For example, when  $g_k$  is the Inner Product function with  $k = \Omega(\log n)$ , the *deterministic* communication complexity of its block-composition with *any*  $f_n$  is asymptotically at most the quantum complexity to the power of 7.

**Keywords:** Communication complexity, quantum information processing, polynomial approximation of Boolean functions, quantum lower bound.

# 1 Introduction and summary of results

Communication complexity studies the inherent communication cost for distributive computations. Let  $F : X \times Y \rightarrow \{0,1\}$  be a function which two parties Alice, who knows  $x \in X$ , and Bob, who knows  $y \in Y$ , wish to compute. The *communication complexity* of  $F$  is the minimum amount of information they need to exchange to compute  $F$  on the worst case input. There are several variants of communication complexity: each of which corresponds to different types of interactions allowed and whether or not small error probability is allowed. For example, we study the following three variants in this paper: deterministic (denoted by  $D(F)$ ), randomized (denoted by  $R(F)$ ), and quantum (denoted by  $Q(F)$ ). In the last two cases, the protocol may err with probability  $\leq 1/3$ . Since its introduction by Yao [37] in 1979, the study of communication complexity has developed into a major branch of complexity theory, with a wide range of applications such as in VLSI design, time-space tradeoff, derandomization, and circuit complexity. The monograph by Kushilevitz and Nisan [23] surveys results up to 1997.

Quantum communication complexity was introduced by Yao [38] in 1993, and has been studied extensively since then. A major problem in this area is to identify problems that have an exponential gap between quantum and classical communication complexities, or to prove that such a problem does not exist.

Exponential gaps were indeed found for several communication tasks ([2, 29, 3, 14, 17, 13]). However, those tasks are either sampling, or computing a partially defined function or a relation. An exponential gap is known for a total Boolean function (checking equality), but in a restricted model that involves a third party (*Simultaneous Message Passing* model) [6]. It remains open to day if super-polynomial gaps are possible for computing a *total Boolean function* in the more commonly studied model of two-party interactive communication. This is one of the most significant problems in quantum communication complexity, and is the question we address in this article.

It is believed that the answer to the above question is “No”:

**Conjecture 1.1** (Log-Equivalence Conjecture). *For any total Boolean function, the quantum (with shared entanglement) and randomized (with shared randomness) communication complexities are polynomially related in the two-way interactive model.*

Besides the lack of a natural candidate for a super-polynomial gap, two other intuitions support this conjecture. The first relates to the approximate version of the well known Log-Rank Conjecture, which states that for any  $F : X \times Y \rightarrow \{0,1\}$ ,  $R(F)$  is polynomially related to  $\widetilde{\text{Logrank}}(F)$ , the logarithm of the smallest rank of a real-valued matrix  $[\tilde{F}(x,y)]_{x,y}$  approximating  $[F(x,y)]_{x \in X, y \in Y}$  entry-wise. It is known that without sharing entanglement, the quantum complexity of  $F$  is at least  $\frac{1}{2}\widetilde{\text{Logrank}}(F)$ . Thus the Log-Equivalence Conjecture follows from the Log-Rank Conjecture, unless there exist exponential gaps between quantum protocols with or without shared entanglement. The existence of such gaps is also a fundamental open problem currently under active investigations.

The second intuition supporting the Log-Equivalence Conjecture is the fact that the similar conjecture is true for the closely related decision tree complexity. Recall that a decision tree algorithm computes a function  $f_n : \{0,1\}^n \rightarrow \{0,1\}$  by making queries of the type “what is the  $i$ ’th bit of the input?” The decision tree complexity of  $f_n$  is the minimum number of queries required to compute  $f_n$  correctly for any input. Making use earlier results of Nisan and Szegedy [27] and Paturi [28], Beals, Buhrman, Cleve, Mosca, and de Wolf [5] proved that the quantum and the deterministic decision tree complexities are polynomially related. This is in sharp contrast with the exponential quantum speedups [34, 35, 10] on *partial functions* achieved by the quantum algorithms of Simon’s and Shor’s.

Razborov's work [31] is a significant progress for the Log-Equivalence Conjecture. He defined the following notion of *symmetric predicates*. Let  $f_n : \{0, 1\}^n \rightarrow \{0, 1\}$  be a symmetric function, i.e.,  $f_n(x)$  depends only on the Hamming weight of  $x$ . A function  $F : \{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}$  is called a *symmetric predicate* if  $F(x, y) = f(x_1 \wedge y_1, x_2 \wedge y_2, \dots, x_n \wedge y_n)$ . The DISJOINTNESS function  $\text{DISJ}_n$  is an important symmetric predicate that has been widely studied:

$$\text{DISJ}_n(x, y) \stackrel{\text{def}}{=} \begin{cases} 1 & \exists i, x_i = y_i = 1, \\ 0 & \text{otherwise.} \end{cases}$$

**Theorem 1.2** (Razborov [31]). *For any symmetric predicate  $F : \{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}$ ,  $D(F) = O(\max\{Q(F)^2, Q(F) \log n\})$ .*

Combined with the  $O(d \log d)$ -bit classical protocol of Huang et al. [16] for deciding if  $x, y \in \{0, 1\}^n$  has Hamming distance  $|x \oplus y| \geq d$ , Razborov's lower bound implies the following.

**Proposition 1.3.** *For any symmetric predicate  $F : \{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\}$ ,  $R(F) = O((Q(F))^2)$ .*

This bound is tight on  $\text{DISJ}_n$ , which admits the largest known quantum-classical gap for total Boolean functions. The class of symmetric predicates is also the most general class of functions on which the Log-Equivalence Conjecture is known to hold.

Notice that Razborov's lower bound method relies on the *symmetry* of  $f_n$ . Thus we aim to develop lower-bound techniques for an arbitrary  $f_n$ , and to derive new quantum lower bounds. To this end, we consider the following class of functions.

**Definition 1.4.** Let  $k, n \geq 1$  be integers. Given  $f_n : \{0, 1\}^n \rightarrow \{0, 1\}$ , and  $g_k : \{0, 1\}^k \times \{0, 1\}^k \rightarrow \{0, 1\}$ , the *block-composition* of  $f_n$  and  $g_k$  is the function  $f_n \square g_k : \{0, 1\}^{nk} \times \{0, 1\}^{nk} \rightarrow \{0, 1\}$  such that on  $x, y \in \{0, 1\}^{nk}$ , with  $x = x_1 x_2 \dots x_n$ , and  $y = y_1 y_2 \dots y_n$ , where  $x_i, y_i \in \{0, 1\}^k$ ,

$$f_n \square g_k(x, y) = f_n(g_k(x_1, y_1), g_k(x_2, y_2), \dots, g_k(x_n, y_n)).$$

Note that a symmetric predicate based on a symmetric  $f_n : \{0, 1\}^n \rightarrow \{0, 1\}$  is the block composition  $f_n \square \wedge$ , where  $\wedge$  denotes the binary AND function. In our Main Lemma, stated and proved in Section 3, we derive a sufficient condition for  $Q(f_n \square g_k)$  to have a strong lower bound. An application of this Main Lemma is the following.

**Theorem 1.5** (Informally). *For any integer  $n \geq 1$  and any function  $f_n : \{0, 1\}^n \rightarrow \{0, 1\}$ , the block composition of  $f_n$  with a  $g_k : \{0, 1\}^k \times \{0, 1\}^k \rightarrow \{0, 1\}$  has polynomially related quantum and randomized communication complexities, if  $g_k$  is sufficiently hard.*

We will define what “sufficiently hard” means precisely. Roughly, it means that  $Q(g_k)$  and  $R(g_k)$  are polynomially related, and some type of discrepancy parameter (Definition 3.2) on  $g_k$  is sufficiently small. We state below an incarnation of the above theorem. Let  $\text{IP}_k : \{0, 1\}^k \times \{0, 1\}^k \rightarrow \{0, 1\}$  be the widely studied Inner Product function

$$\text{IP}_k(x, y) \stackrel{\text{def}}{=} \sum_i x_i y_i \pmod{2}, \quad \forall x, y \in \{0, 1\}^k.$$

**Corollary 1.6.** *For any integers  $k$  and  $n$  with  $k \geq 2 \log_2 n + 5$ , and for an arbitrary  $f_n : \{0, 1\}^n \rightarrow \{0, 1\}$ ,  $D(f_n \square \text{IP}_k) = O((Q(f_n \square \text{IP}_k))^7)$ .*

The above corollary also holds for a random  $g_k$  with high probability. Our technique can also be applied to symmetric predicates, thus giving an alternative proof to Razborov’s result, albeit with a weaker parameter.

**Theorem 1.7.** *For any symmetric  $f_n : \{0, 1\}^n \rightarrow \{0, 1\}$ ,  $R(f_n \square \wedge) = O((Q(f_n \square \wedge))^3)$ .*

Our approach is inspired by how the Log-Equivalence result in decision tree complexity was proved: for any  $f_n : \{0, 1\}^n \rightarrow \{0, 1\}$ , both the quantum and the deterministic decision tree complexities were shown [27, 5] to be polynomially related to the *approximate polynomial degree* (or, approximate degree for short)  $\widetilde{\deg}(f_n)$ , which is the smallest degree of a real polynomial that approximate  $f_n$  to be within  $1/3$  on any  $0/1$  inputs. In our Main Lemma, we derive a sufficient condition on  $n$  and  $k$ , and  $g_k$  such that  $Q(f_n \square g_k) = \Omega(\widetilde{\deg}(f_n))$ , for any  $f_n$ . The randomized upper bound is obtained by simulating a decision tree algorithm for  $f_n$ , and whenever one input bit of  $f_n$  is needed, the protocol calls a sub-protocol for computing  $g_k$  on the corresponding block. Under some hardness assumption on  $g_k$ , those upper and lower bounds are polynomially related.

The approach for proving a quantum lower bound using an approximate degree lower bound is termed the *polynomial method* in the literature of quantum decision trees. Razborov’s lower bound on DISJ can be viewed an application of the polynomial method as well. This is because, he showed that if there is a  $q$ -qubit protocol for DISJ $_n$ , then there is a  $O(q)$ -degree polynomial approximating OR $_n$ . Thus the quantum lower bound of  $\Omega(\sqrt{n})$  follows from the same lower bound on  $\widetilde{\deg}(\text{OR}_n)$  due to Nisan and Szegedy [27] and Paturi [28]. We emphasize this connection of approximating polynomial and quantum protocol is not obvious at all and it makes use the symmetry of DISJ critically.

We avoid the dependence of Razborov’s proof on the symmetry property of  $f_n$  by taking the *dual* approach of the polynomial method. We show that from the linear programming formulation of polynomial approximation, we can obtain a “witness” for  $f_n$  requiring a high approximate degree. This witness is then turned into a “witness” for the hardness of  $f_n \square g_k$ , under certain assumptions. While the approximate polynomial degree has been used to prove lower bounds, and its dual formulation has been known to several researchers [30, 36], our application of the dual form appears to be the first demonstration of its usefulness in proving new results. We note that there are several recent works that use the duality of linear (or semidefinite) programming in the context of communication complexity [25, 32, 24, 26]. Those applications of duality, however, do not involve the type of polynomial approximation of Boolean functions considered here.

Before we proceed to the proofs, we briefly review some other closely related works. Buhrman and de Wolf [8] are probably the first to systematically study the relationship of polynomial representations and communication complexity. However, their result applies to error-free quantum protocols, while we consider bounded-error case. Klauck [18] proved strong lower bounds for some symmetric predicates such as MAJORITY based on the properties of their Fourier coefficients. The same author formulated a lower bound framework that includes several known lower bound methods [19]. It would be interesting to investigate the limitations of our polynomial method in this framework. After preparing this draft, we learned about an independent work by Sherstov [33], who used a similar approach to prove similar results. We will compare our work with his in the concluding section.

## 2 Preliminaries

### 2.1 Communication complexities and quantum lower bound by approximate trace norm

Denote the domain of a function by  $\text{dom}(\cdot)$ . For a positive integer  $n$ , denote by  $\mathcal{F}_n \stackrel{\text{def}}{=} \{f_n : \{0,1\}^n \rightarrow \{0,1\}\}$ , and by  $\mathcal{G}_k \stackrel{\text{def}}{=} \{g_k : \text{dom}(g_k) \rightarrow \{0,1\}, \text{dom}(g_k) \subseteq \{0,1\}^k \times \{0,1\}^k\}$ . For the rest of this article  $f_n \in \mathcal{F}_n$  and  $g_k \in \mathcal{G}_k$ , for some integers  $n, k \geq 1$ .

If  $F \in \mathcal{G}_n$  is a total function, we also denote by  $F$  the  $\{0,1\}^{2^n \times 2^n}$  matrix  $[F(x,y)]_{x,y \in \{0,1\}^n}$ . Consider the computation of  $F \in \mathcal{G}_n$  on  $(x,y) \in \text{dom}(F)$  when the input  $x$  is known to a party Alice and  $y$  is known to another party Bob. Unless  $F(x,y)$  trivially depends only on  $x$  or  $y$ , Alice and Bob will have to communicate before they could determine  $F(x,y)$ . The worst case cost of communication is called the communication complexity of  $F$ .

The communication complexity depends on the information processing power of Alice and Bob, and the requirement on the accuracy of the outcome of a protocol. Thus we have various communication complexities: deterministic (denoted by  $D(f)$ ), randomized ( $R_\epsilon(f)$ ), and quantum ( $Q_\epsilon(f)$ ), when the protocols are restricted to be deterministic, randomized, and quantum, respectively, and  $\epsilon \in (0, 1/2)$  is a constant that upper-bounds the error probability of the protocols. In the randomized and the quantum cases we allow Alice and Bob share unlimited amount of randomness or quantum entanglement, respectively. Different choices of  $\epsilon$  only result in a change of a constant factor in the communication complexities, by a standard application of the Chernoff Bound. Thus we may omit the subscripts in  $R_\epsilon(F)$  and  $Q_\epsilon(F)$  for asymptotic estimations.

A powerful method for proving quantum communication complexity lower bounds is the following lemma, which was obtained by Razborov [31], extending a lemma of Yao [38]. Recall that the trace norm of a matrix  $A \in \mathbb{R}^{N \times M}$  is  $\|A\|_{tr} \stackrel{\text{def}}{=} \text{trace} \sqrt{A^T A} = \text{trace} \sqrt{A A^T}$ . Let  $F$  be a partial Boolean function defined on a subset  $\text{dom}(F) \subseteq X \times Y$ . The *approximate trace norm* of  $F$  with error  $\epsilon$ ,  $0 \leq \epsilon < 1/2$ , is

$$\|F\|_{\epsilon, \text{tr}} \stackrel{\text{def}}{=} \min\{\|\tilde{F}\|_{tr} : \tilde{F} \in \mathbb{R}^{N \times M}, \forall (x,y) \in \text{dom}(F), |\tilde{F}(x,y) - F(x,y)| \leq \epsilon\}.$$

**Lemma 2.1** (Razborov-Yao[31, 38]). *For any partial Boolean function  $F$  whose domain is a subset of  $X \times Y$ ,  $Q_\epsilon(F) = \Omega(\log \frac{\|F\|_{\epsilon, \text{tr}}}{\sqrt{|X| \cdot |Y|}})$ .*

For matrix  $B$ , denote by  $\|B\|$  its operator norm. Since for any matrix  $A$ ,  $\|A\|_{tr} = \sup_{B, \|B\|=1} |\text{trace}(B^T A)|$ , we have

$$\|A\|_{tr} \geq \frac{|\text{trace}(B^T A)|}{\|B\|}, \quad \forall B \neq 0.$$

Therefore, in order to prove that  $\|A\|_{tr}$  is large, we need only to find a  $B$  so that  $|\text{trace}(B^T A)|/\|B\|$  is large.

### 2.2 Approximate polynomial degree

The study of low degree polynomial approximations of Boolean function under the  $\ell_\infty$  norm was pioneered by Nisan and Szegedy [27] and Paturi [28], and has since then been a powerful tool in studying concrete complexities, including the quantum decision tree complexity (c.f. the survey by Buhrman and de Wolf [9]).

Let  $f \in \mathcal{F}_n$ . A real polynomial  $\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R}$  is said to *approximate*  $f$  with an error  $\epsilon$ ,  $0 < \epsilon < 1/2$ , if

$$|f(x) - \tilde{f}(x)| \leq \epsilon, \quad \forall x \in \{0, 1\}^n.$$

The *approximate degree* of  $f$ , denoted by  $\widetilde{\deg}_\epsilon(f)$  is smallest degree of a polynomial approximating  $f$  with an error  $\epsilon$ . Difference choices for  $\epsilon$  only result in a constant factor difference in the approximate degrees. Thus we omit the subscript  $\epsilon$  for asymptotic estimations.

While the approximate degree of symmetric functions has a simple characterization [27, 28], it is difficult to determine in general. For example, the approximate degree of the two level AND-OR trees is still unknown. On the other hand,  $\widetilde{\deg}(f)$  is polynomially related to the deterministic decision tree complexities  $T(f)$ . Formally,  $T(f)$  is defined to be the minimum integer  $k$  such that there is an ordered full binary tree  $T$  of depth  $k$  satisfying the following properties: (a) each non-leaf vertex is labeled by a variable  $x_i$ , and each leaf is labeled by either 0 or 1 (but not both); (b) for any  $x \in \{0, 1\}^n$ , the following walk leads to a leaf labeled with  $f(x)$ : start from the root, at each non-leaf vertex labeled with  $x_i$ , take the left edge if  $x_i = 0$ , and take the right edge otherwise.

**Theorem 2.2** (Nisan and Szegedy [27], Beals et al. [5]). *For any Boolean function  $f_n$ , there are constants  $c_1$  and  $c_2$  such that  $c_1 T^{1/6}(f) \leq \widetilde{\deg}(f) \leq c_2 T(f)$ .*

The exponent  $1/6$  is not known to be optimal. The conjectured value is  $1/2$ .

As observed by Buhrman, Cleve, and Wigderson [7], a decision tree algorithm can be turned into a communication protocol for a related problem. In such a protocol for  $f_n \square g_k$ , one party simulates the decision tree algorithm for  $f_n$ , and initiates a sub-protocol for computing  $g_k$  whenever one input bit of  $f_n$  is needed. The sub-protocol repeats an optimal protocol for  $g_k$  for  $O(\log \widetilde{\deg}(f_n))$  times, ensuring that the error probability is  $\leq \frac{1}{3(\deg(f_n)/c_1)^6}$ . Thus the larger protocol computes  $f_n \square g_k$  with error probability  $\leq 1/3$ , and exchanges  $O(R(g_k) \widetilde{\deg}^6(f_n) \log \widetilde{\deg}(f_n))$  bits.

**Proposition 2.3** ([7, 5]). *For any function  $f_n \in \mathcal{F}_n$  with  $\widetilde{\deg}(f_n) = d$ , and any  $g_k \in \mathcal{G}_k$ ,  $R(f_n \square g_k) = O(R(g_k) d^6 \log d)$ .*

### 3 The Main Lemma

In this section, we prove that under some assumptions,  $Q(f_n \square g_k) = \Omega(\widetilde{\deg}(f_n))$ . This is shown by turning a “witness” for  $f_n$  requiring a high approximate degree into a “witness” for the hardness of  $f_n \square g_k$ .

#### 3.1 Witness of high approximate degree

We now fix a function  $f_n \in \mathcal{F}_n$  with  $\widetilde{\deg}_\epsilon(f_n) = d$ . For  $w \in \{0, 1\}^n$ , denote by  $\chi_w \in \mathcal{F}_n$  the function  $\chi_w(x) = (-1)^{w \cdot x}$ . Then there is no feasible solution to the following linear system, where the unknowns are  $\alpha_w$ :

$$-\epsilon + f(x) \leq \sum_{w: |w| < d} (-1)^{w \cdot x} \alpha_w \leq f(x) + \epsilon, \quad \forall x \in \{0, 1\}^n. \quad (1)$$

By the duality of linear programming, there exist  $q_x^+ \geq 0$  and  $q_x^- \geq 0$ ,  $x \in \{0, 1\}^n$ , such that

$$\sum_x (q_x^+ - q_x^-) \cdot \chi_w = 0, \quad \forall w, |w| < d, \quad \text{and,}$$

$$\sum_x (q_x^+ - q_x^-) f(x) + \epsilon(q_x^+ + q_x^-) < 0. \quad (2)$$

Define  $q : \{0, 1\}^n \rightarrow \mathbb{R}$  as  $q(x) = q_x^- - q_x^+$ . Then

$$q^T \chi_w = 0, \quad \text{and,} \quad \|q\|_1 < \frac{1}{\epsilon} q^T f.$$

Without loss of generality, assume that  $q^T f = 1$  (otherwise this will hold after multiplying  $q$  with an appropriate positive number). Then  $\|q\|_1 < 1/\epsilon$ .

Since  $q$  is orthogonal to all polynomials of degree less than  $d$ , it has non-zero Fourier coefficients only on higher frequencies:  $q = \sum_{w: |w| \geq d} \hat{q}_w \chi_w$ , where  $\hat{q}_w = \frac{1}{2^n} \sum_x q(x) \chi_w(x)$ . Since  $\|q\|_1 < 1/\epsilon$ , those Fourier coefficients must be small:  $|\hat{q}_w| < \frac{1}{2^n \epsilon}$ ,  $\forall w : |w| \geq d$ .

We summarize the above discussion in the following lemma.

**Lemma 3.1.** *Let  $\epsilon \in \mathbb{R}$ ,  $0 \leq \epsilon < 1/2$ . For any  $f \in \mathcal{F}_n$ , there exists a function  $q : \{0, 1\}^n \rightarrow \mathbb{R}$  such that: (a)  $q^T f = 1$ , (b)  $\|q\|_1 < 1/\epsilon$ , (c)  $|\hat{q}_w| \leq \frac{1}{2^n \epsilon}$ , for all  $w \in \{0, 1\}^n$ , and (d)  $\hat{q}_w = 0$  whenever  $|w| < \widetilde{\deg}_\epsilon(f_n)$ .*

### 3.2 Witness of large approximate trace norm

In order to convert a witness of high approximate degree for  $f_n$  to that of large approximate trace norm for  $f_n \square g_k$ , we need to require that  $g_k$  satisfies certain property, which we now formulate. Let  $I_A, I_B \subseteq \{0, 1\}^k$ . For  $b \in \{0, 1\}$ , we identify a probability distribution  $\mu$  on  $g_k^{-1}(b) \cap I_A \times I_B$  with its representation as a matrix in  $\mathbb{R}^{I_A \times I_B}$ , and call it a  $b$ -distribution.

Recall that the *discrepancy* of  $g_k \in \mathcal{G}_k$ , denoted by  $\text{disc}(g_k)$ , is

$$\text{disc}(g_k) \stackrel{\text{def}}{=} \min_{\mu} \max_{I_A, I_B \subseteq \{0, 1\}^k} \left| \sum_{(x, y) \in I_A \times I_B} \mu(x, y) (-1)^{g_k(x, y)} \right|,$$

where  $\mu$  ranges over all distributions on  $\text{dom}(g_k)$ . We define a more restricted concept of discrepancy.

**Definition 3.2.** The *spectral discrepancy* of  $g_k \in \mathcal{G}_k$ , denoted by  $\rho(g_k)$ , is the minimum  $r \in \mathbb{R}$  such that there exist  $I_A, I_B \subseteq \{0, 1\}^k$ , and  $b$ -distributions  $\mu_b \in \mathbb{R}^{I_A \times I_B}$  for  $g_k$ ,  $b \in \{0, 1\}$ , satisfying the following conditions: (1)  $\sqrt{|I_A| \cdot |I_B|} \cdot \|\frac{\mu_0 + \mu_1}{2}\| \leq 1 + r$ , and, (2)  $\sqrt{|I_A| \cdot |I_B|} \cdot \|\frac{\mu_0 - \mu_1}{2}\| \leq r$ .

While (1) appears contrived, it will only be used in deriving a general lower bound on quantum communication complexity. In all of explicit applications, (1) is trivially satisfied with  $\|\frac{\mu_0 + \mu_1}{2}\| = 1/\sqrt{|I_A| \cdot |I_B|}$ .

Kremer [21] showed that  $\log(1/\text{disc}(g_k))$  is a lower bound for the quantum communication complexity of  $g_k$  when no shared entanglement is allowed. Linial and Shraibman [25] recently showed that the lower bound holds even when shared entanglement is allowed.

**Theorem 3.3** (Linial and Shraibman [25]). *For any  $g_k \in \mathcal{G}_k$ ,  $Q(g_k) = \Omega(\log(1/\text{disc}(g_k)))$ .*

Suppose that  $\rho(g_k)$  is achieved with  $I_A, I_B$  and  $\mu$ . Since for any  $I'_A \subseteq I_A, I'_B \subseteq I_B$ ,

$$\left| \sum_{(x, y) \in I'_A \times I'_B} \mu(x, y) (-1)^{g_k(x, y)} \right| \leq \sqrt{|I'_A| \cdot |I'_B|} \cdot \|\frac{\mu_0 - \mu_1}{2}\| \leq \frac{\sqrt{|I'_A| \cdot |I'_B|}}{\sqrt{|I_A| \cdot |I_B|}} \rho(g_k) \leq \rho(g_k),$$

we have

$$\begin{aligned}
\text{disc}(g_k) &\leq \max_{J_A, J_B \subseteq \{0,1\}^k} \left| \sum_{(x,y) \in J_A \times J_B} \mu(x,y) (-1)^{g_k(x,y)} \right| \\
&\leq \max_{I'_A \subseteq I_A, I'_B \subseteq I_B} \left| \sum_{(x,y) \in I'_A \times I'_B} \mu(x,y) (-1)^{g_k(x,y)} \right| \\
&\leq \rho(g_k).
\end{aligned}$$

It follows from Theorem 3.3,

**Proposition 3.4.** *For any  $g_k \in \mathcal{G}_k$ ,  $Q(g_k) = \Omega(\log \frac{1}{\rho(g_k)})$ .*

With the concept of spectral discrepancy, we are now ready to state and prove our Main Lemma.

**Lemma 3.5** (Main Lemma). *Let  $n, k \geq 1$  be integers,  $g_k \in \mathcal{G}_k$ , and  $f_n \in \mathcal{F}_n$ . If  $\rho(g_k) \leq \frac{\widetilde{\text{deg}}(f_n)}{2\epsilon n}$ , then  $Q(f_n \square g_k) = \Omega(\widetilde{\text{deg}}(f_n))$ .*

*Proof.* Let  $d \stackrel{\text{def}}{=} \widetilde{\text{deg}}(f_n)$ , and  $F \stackrel{\text{def}}{=} f_n \square g_k$ . Suppose  $\rho \stackrel{\text{def}}{=} \rho(g_k)$  is achieved with  $I_A, I_B \subseteq \{0,1\}^k$ , and  $\mu_b$ ,  $b \in \{0,1\}$ . Denote  $K_A \stackrel{\text{def}}{=} |I_A|$ ,  $K_B \stackrel{\text{def}}{=} |I_B|$ . Let  $F_1$  be the restriction of  $f_n \square g_k$  on  $(I_A \times I_B)^{\otimes n} \cap \text{dom}(F)$ . We shall prove the desired lower bound on  $F_1$ . By Lemma 2.1, it suffices to prove a lower bound on  $\|F_1\|_{\epsilon', \text{tr}}$  for  $\epsilon' = 1/6$ . Let  $q$  be the function that exists by Lemma 3.1 with respect to  $f_n$  and  $\epsilon = 1/3$ .

For a partition  $\{w_1, w_2, \dots, w_K\}$  of  $[n]$ , and matrices  $A_1, A_2, \dots, A_k \in K_A \times K_B$ , denote by  $\bigotimes_{k=1}^K A_k^{w_k} \in (\mathbb{R}^{K_A \times K_B})^{\otimes n}$  the product element that has  $A_k$  in the components indexed by  $w_k$ . Denote by  $\bar{w}$  the complement of  $w$ . Define  $h \in (\mathbb{R}^{K_A \times K_B})^{\otimes n}$  as follows

$$h \stackrel{\text{def}}{=} \sum_{z \in \{0,1\}^n} q(z) \cdot \bigotimes_{i=1}^n \mu_{z_i}^{\otimes \{i\}}. \quad (3)$$

For a matrix  $A = [A_{ij}]$ , denote by  $\|A\|_1 \stackrel{\text{def}}{=} \sum_{i,j} |A_{ij}|$ . Then  $\|\mu_0\|_1 = \|\mu_1\|_1 = 1$ , and for any  $z \in \{0,1\}^n$ ,

$$\left\| \bigotimes_{i=1}^n \mu_{z_i}^{\otimes \{i\}} \right\|_1 = \prod_{i=1}^n \|\mu_{z_i}\|_1 = 1.$$

Since for a different  $z$ , the set of the non-zero entries in  $\bigotimes_{i=1}^n \mu_{z_i}^{\otimes \{i\}}$  is disjoint,

$$\|h\|_1 = \sum_{z \in \{0,1\}^n} |q(z)| \left\| \bigotimes_{i=1}^n \mu_{z_i}^{\otimes \{i\}} \right\|_1 = \|q\|_1 \leq 1/\epsilon.$$

Note that  $\text{tr}((\bigotimes_{i=1}^n \mu_{z_i}^{\otimes \{i\}})^T F) = f(z_1, z_2, \dots, z_n)$ . Thus

$$\text{tr}(h^T F) = q^T f_n = 1.$$

Fix an  $\tilde{F} \in (\mathbb{R}^{K_A \times K_B})^{\otimes n}$  with  $|F_1(x, y) - \tilde{F}(x, y)| \leq \epsilon'$ ,  $\forall (x, y) \in \text{dom}(F_1)$ . Then,

$$|\text{tr}(h^T \tilde{F})| = \left| \sum_{(x,y) \in \text{dom}(F_1)} h(x, y) \tilde{F}(x, y) \right| \geq \left| \sum_{(x,y) \in \text{dom}(F)} h(x, y) F(x, y) \right| - \epsilon' \|h\|_1 \geq 1 - \epsilon'/\epsilon \geq 1/2.$$



Therefore,

$$\|\tilde{F}\|_{\text{tr}} \geq \frac{|\text{tr}(h^T \tilde{F})|}{\|h\|} \geq \frac{1}{2\|h\|}. \quad (4)$$

Hence we need only to prove that  $\|h\|$  is very small. To this end we first express  $h$  using the Fourier representation of  $q$ :

$$\begin{aligned} h &= \sum_{z \in \{0,1\}^n} \sum_{w: |w| \geq d} \hat{q}_w (-1)^{w \cdot z} \cdot \bigotimes_{i=1}^n \mu_{z_i}^{\otimes \{i\}} \\ &= \sum_{w: |w| \geq d} \hat{q}_w \cdot \sum_{z \in \{0,1\}^n} (-1)^{w \cdot z} \cdot \bigotimes_{i=1}^n \mu_{z_i}^{\otimes \{i\}} \\ &= \sum_{w: |w| \geq d} \hat{q}_w \cdot ((\mu_0 + \mu_1)^{\otimes \bar{w}}) \otimes ((\mu_0 - \mu_1)^{\otimes w}). \end{aligned} \quad (5)$$

Using  $\hat{q}_w \leq 1/\epsilon 2^n$ ,

$$\|h\| \leq \sum_{w: |w| \geq d} |\hat{q}_w| \|\mu_0 + \mu_1\|^{n-|w|} \cdot \|\mu_0 - \mu_1\|^{|w|} \leq \frac{1}{\epsilon} \sum_{\ell, \ell \geq d} \binom{n}{\ell} \cdot \left\| \frac{\mu_0 + \mu_1}{2} \right\|^{n-|w|} \cdot \left\| \frac{\mu_0 - \mu_1}{2} \right\|^{|w|}. \quad (6)$$

By the choice of  $\mu_0$  and  $\mu_1$ ,  $\left\| \frac{\mu_0 + \mu_1}{2} \right\| \leq \frac{1+\rho}{\sqrt{K_A K_B}}$ , and  $\left\| \frac{\mu_0 - \mu_1}{2} \right\| \leq \frac{\rho}{\sqrt{K_A K_B}}$ . Thus

$$\|h\| \leq \frac{(1+\rho)^n}{\epsilon (K_A K_B)^{n/2}} \sum_{\ell: \ell \geq d} \binom{n}{\ell} \rho^\ell. \quad (7)$$

If  $\rho \leq \frac{d}{2en}$ , using  $\binom{n}{\ell} \leq \left(\frac{en}{\ell}\right)^\ell$ , and  $(1+\rho)^n \leq e^{\rho n}$ , we have

$$\|h\| \leq \frac{e^{\rho n}}{\epsilon (K_A K_B)^{n/2}} \sum_{\ell \geq d} \left(\frac{en\rho}{\ell}\right)^\ell \leq \frac{e^{\rho n}}{\epsilon (K_A K_B)^{n/2}} \sum_{\ell \geq d} \left(\frac{d}{2\ell}\right)^\ell \leq \frac{2}{\epsilon (K_A K_B)^{n/2}} e^{-.5d}.$$

Together with Equation 4, this implies  $\|\tilde{F}\|_{\text{tr}} \geq \frac{\epsilon}{4} \cdot (K_A K_B)^{n/2} \cdot e^{.5d}$ . Thus  $\|F_1\|_{1/6, \text{tr}} \geq \frac{1}{24} \cdot (K_A K_B)^{n/2} \cdot e^{.5d}$ . Plugging this inequality to the Razborov-Yao Lemma, we have  $Q(F) \geq Q(F_1) = \Omega(d)$ .  $\square$

## 4 Applications

We now apply the Main Lemma to derive two quantum lower bounds. The first deals with those  $g_k$  that have polynomially related quantum and randomized communication complexities. As a concrete example we consider  $g_k$  being the INNER PRODUCT function. The second result shows that without this knowledge on  $g_k$ , we may still able to obtain strong quantum lower bounds. This is done through a “hardness amplification” technique that makes use of the self-similarity of the function considered. We demonstrate this technique by proving Theorem 1.7.

## 4.1 Composition with hard $g_k$

We now restate Theorem 1.5 rigorously.

**Theorem 4.1.** *Let  $n, k \geq 1$  be integers and  $g_k \in \mathcal{G}_k$ . If  $Q(g_k)$  and  $R(g_k)$  are polynomially related, so is  $Q(f_n \square g_k)$  and  $R(f_n \square g_k)$  for any  $f_n \in \mathcal{F}_n$  and for  $\rho(g_k) \leq \frac{1}{2en}$ .*

*Proof.* If  $f_n$  or  $g_k$  is a constant function,  $Q(f_n \square g_k) = R(f_n \square g_k) = 0$ , hence the statement holds. Otherwise, one can fix the value of all but one input block so that  $f_n \square g_k$  computes  $g_k$  on the remaining block. Thus  $Q(f_n \square g_k) \geq Q(g_k)$ . By Main Lemma, under the assumption that  $\rho(g_k) \leq \frac{1}{2en}$ ,  $Q(f_n \square g_k) = \Omega(\widetilde{\deg}(f_n))$ . Thus  $Q(f_n \square g_k) = \Omega(\widetilde{\deg}(f_n) + Q(g_k))$ . On the other hand  $R(f_n \square g_k) = O(R(g_k) \widetilde{\deg}^6(f_n) \log \widetilde{\deg}(f_n))$ , by Proposition 2.3. Thus, under the assumption that  $R(g_k)$  and  $Q(g_k)$  are polynomially related, so are  $Q(f_n \square g_k)$  and  $R(f_n \square g_k)$ .  $\square$

Similarly, the same statement holds with  $R(f_n \square g_k)$  and  $R(g_k)$  replaced by  $D(f_n \square g_k)$  and  $D(g_k)$ , respectively. Estimating  $\rho(g_k)$  is unfortunately difficult in general. However, if we can show  $\rho(g_k) = \exp(-\Omega(k^c))$  for some constant  $c$ , it implies  $R(g_k)$  and  $Q(g_k)$  are polynomially related, by Proposition 3.4. Thus  $Q(f_n \square g_k)$  and  $R(f_n \square g_k)$  are polynomially related for  $k \geq \log_2^{1/c}(2en)$ .

We now prove Corollary 1.6.

*Proof of Corollary 1.6.* We need only to consider the case that  $f_n$  is not a constant function. Then  $Q(f_n \square g_k) = \Omega(\text{IP}_k)$ . It is known that  $Q(\text{IP}_k) = \Omega(k)$  [11]. Thus  $Q(f_n \square g_k) = \Omega(k)$ . Let  $K \stackrel{\text{def}}{=} 2^k$ ,  $I_A \stackrel{\text{def}}{=} \{0, 1\}^k - \{0^k\}$ , and  $I_B \stackrel{\text{def}}{=} \{0, 1\}^k$ . For  $b \in \{0, 1\}$ , let  $\mu_b$  be the uniform distribution on  $\{(x, y) : \text{IP}(x, y) = b, x \neq 0\}$ . Then

$$\left\| \frac{\mu_0 + \mu_1}{2} \right\| = 1/\sqrt{K(K-1)}, \quad \text{and} \quad \left\| \frac{\mu_0 - \mu_1}{2} \right\| = 1/((K-1)\sqrt{K}).$$

Thus  $\rho(\text{IP}_k) \leq 1/\sqrt{K-1}$ . When  $k \geq 2\log_2 n + 5 > \log_2(4e^2 n^2 + 1)$ , we have  $\rho(\text{IP}_k) \leq 1/2en \leq \widetilde{\deg}(f_n)/(2en)$ . By Main Lemma 3.5, this implies  $Q(f_n \square \text{IP}_k) = \Omega(\widetilde{\deg}(f_n))$ . Therefore,  $Q(f_n \square \text{IP}_k) = \Omega(k + \widetilde{\deg}(f_n))$ .

On the other hand,  $D(f_n \square \text{IP}_k) \leq kT(f_n)$ , which is  $O(k \widetilde{\deg}^6(f_n))$  by Theorem 2.2. Thus  $D(f_n \square \text{IP}_k) = O(Q^7(f_n \square \text{IP}_k))$ .  $\square$

We remark that since for a random  $g_k$ ,  $\rho(g_k) = \exp(-\Omega(k))$ , the above corollary holds for most  $g_k$  up to a constant additive difference in the bound for  $k$ .

## 4.2 Composition with Set Disjointness

In this section we prove Theorem 1.7. We introduce some notions following [31]. For an integer  $k \geq 1$ , let  $[k] \stackrel{\text{def}}{=} \{1, 2, \dots, n\}$ . For an integer  $p$ ,  $0 \leq p \leq k$ , denote by  $[k]^p$  the set of  $p$ -element subsets of  $[k]$ . For integers  $s$  and  $p$  with  $0 \leq s \leq p \leq k/2$ , denote by  $J_{k,p,s} \in \{0, 1\}^{[k]^p \times [k]^p}$  the indicator function for  $|x \cap y| = s$ . That is, for any  $(x, y) \in [k]^p \times [k]^p$ ,

$$(J_{k,p,s})_{x,y} \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } |x \cap y| = s, \\ 0 & \text{otherwise.} \end{cases}$$

The spectrum of these combinatorial matrices are described by *Hahn polynomials* [12]. We will use a formula given by Knuth [20].

**Proposition 4.2** (Knuth). *Let  $p \leq k/2$ . Then the matrices  $J_{k,p,s}$ ,  $0 \leq s \leq p$ , share the same eigenspaces  $E_0, E_1, \dots, E_p$ , and the eigenvalue corresponding to the eigenspace  $E_t$ ,  $0 \leq t \leq p$ , is given by*

$$\sum_{i=\max\{0,s+t-p\}}^{\min\{s,t\}} (-1)^{t-i} \binom{t}{i} \binom{p-i}{s-i} \binom{k-p-t+i}{p-s-t+i}. \quad (8)$$

We actually need only to consider  $s \in \{0, 1\}$ . Effectively, we are restricting  $\text{DISJ}_k$  on  $\{(u, v) : u, v \in [k]^p, |u \cap v| \leq 1\}$ . Denote this restriction by  $\text{DISJ}_k^{\leq 1}$ .

**Lemma 4.3.** *Let  $n, k \geq 1$  be integers,  $f_n \in \mathcal{F}_n$ , and  $k \geq \frac{6en}{\deg(f_n)}$ . Then  $Q(f_n \square \text{DISJ}_k^{\leq 1}) = \Omega(\widetilde{\deg}(f_n))$ .*

*Proof.* Let  $p \stackrel{\text{def}}{=} k/3$  and  $M \stackrel{\text{def}}{=} \binom{k}{p}$ . Let  $w_s \stackrel{\text{def}}{=} |(\text{DISJ}_k^{\leq 1})^{-1}(s)|$ ,  $s \in \{0, 1\}$ . That is,

$$w_0 = \binom{k}{p} \binom{k-p}{p} = M \binom{k-p}{p}, \quad \text{and}, \quad w_1 = \binom{k}{p} \binom{p}{1} \binom{k-p}{p-1} = M \binom{p}{1} \binom{k-p}{p-1}.$$

Let  $\mu_s$ ,  $s \in \{0, 1\}$ , be the distribution matrix for the uniform distribution on the  $s$ -inputs of  $\text{DISJ}_k^{\leq 1}$ . That is,

$$\mu_0 \stackrel{\text{def}}{=} \frac{1}{w_0} J_{k,p,0}, \quad \text{and}, \quad \mu_1 \stackrel{\text{def}}{=} \frac{1}{w_1} J_{k,p,1}.$$

By Proposition 4.2,  $\mu_0$  and  $\mu_1$  have the same eigenspaces. Furthermore, if  $\lambda_{s,t}$ ,  $s \in \{0, 1\}$  and  $0 \leq t \leq p$ , is the eigenvalue of  $\mu_s$  for the eigenspace  $E_t$ ,

$$\lambda_{s,t} = \frac{1}{w_s} \sum_{i=\max\{0,s+t-p\}}^{\min\{s,t\}} (-1)^{t-i} \binom{t}{i} \binom{p-i}{s-i} \binom{k-p-t+i}{p-s-t+i}, \quad (9)$$

and

$$\|\mu_0 - \mu_1\| = \max_{t: 0 \leq t \leq p} |\lambda_{0,t} - \lambda_{1,t}|. \quad (10)$$

After simplification,

$$\lambda_{0,t} = \frac{(-1)^t}{M} \frac{\binom{k-p-t}{p-t}}{\binom{k-p}{p}}, \quad \text{and}, \quad \lambda_{1,t} = \frac{(-1)^t}{M} \left( \frac{\binom{k-p-t}{p-1-t}}{\binom{k-p}{p-1}} - \frac{t \binom{k-p-t+1}{p-1-t+1}}{p \binom{k-p}{p-1}} \right).$$

Since  $\lambda_{0,0} = \lambda_{1,0} = 1$ , we only need to bound  $\max_t |\lambda_{0,t} - \lambda_{1,t}|$  for  $t \geq 1$ . From Proposition 4.2,

$$\begin{aligned} \lambda_{0,t} - \lambda_{1,t} &= \frac{(-1)^t}{M} \frac{\binom{k-p-t}{p-t}}{\binom{k-p}{p}} \left( 1 - \frac{p-t}{p} + \frac{t(k-p-t+1)}{p^2} \right) \\ &= (-1)^t \frac{1}{M} \frac{\binom{k-p-t}{p-t}}{\binom{k-p}{p}} \frac{t(k-t+1)}{p^2}. \end{aligned}$$

With  $k = 3p$ ,

$$\frac{t \binom{k-p-t}{p-t}}{\binom{k-p}{p}} = \frac{t \cdot p \cdot (p-1) \dots (p-t+1)}{(k-p) \cdot (k-p-1) \dots (k-p-t+1)} \leq \left( \frac{p}{k-p} \right)^t \cdot t = \frac{1}{2^t} t \leq \frac{1}{2}.$$

Hence

$$|\lambda_{0,t} - \lambda_{1,t}| \leq \frac{1}{2} \cdot \frac{k-t+1}{Mp^2} \leq \frac{1}{2} \cdot \frac{k}{M(\frac{k}{3})^2} = \frac{6}{Mk}. \quad (11)$$

Therefore,  $M\|\frac{\mu_0+\mu_1}{2}\| \leq \frac{3}{k}$ . Since  $\frac{\mu_0+\mu_1}{2}$  is doubly stochastic,  $\|\frac{\mu_0+\mu_1}{2}\| = 1$ . Thus we have

$$\rho(g_k) \leq 3/k. \quad (12)$$

Therefore, when  $k \geq 6en/d$ , we have  $\rho(g_k) \leq d/(2en)$ . By Main Lemma 3.5, this implies  $Q(f_n \square \text{DISJ}_k^{\leq 1}) = \Omega(\widetilde{\deg}(f_n))$ .  $\square$

Let  $f_n \in \mathcal{F}_n$  be a symmetric function. Following [31], define

$$\ell_0(f_n) \stackrel{\text{def}}{=} \max\{m : 1 \leq m \leq n/2, f_n(1^m 0^{n-m}) \neq f_n(1^{m-1} 0^{n-m+1})\} \cup \{0\},$$

and

$$\ell_1(f_n) \stackrel{\text{def}}{=} \max\{n-m : n/2 \leq m < n, f_n(1^m 0^{n-m}) \neq f_n(1^{m+1} 0^{n-m-1})\} \cup \{0\}.$$

We will use the following result in proving quantum lower bounds on  $f_n \square \wedge$ .

**Theorem 4.4** (Paturi [28]). *Let  $f_n \in \mathcal{F}_n$  be symmetric. Then for some universal constant  $c$ ,  $\widetilde{\deg}(f_n) \geq c\sqrt{n(\ell_0(f_n) + \ell_1(f_n))}$ .*

**Theorem 4.5.** *For any symmetric  $f_n \in \mathcal{F}_n$ ,  $Q(f_n \square \wedge) = \Omega(n^{1/3} \ell_0^{2/3}(f_n) + \ell_1(f_n))$ .*

The lower bound is weaker than Razborov's, which is

$$Q((f_n \square \wedge)) = \Omega(\sqrt{n\ell_0(f_n)} + \ell_1(f_n)). \quad (13)$$

In the following proof, we first show that  $Q(f_n \square \wedge) = \Omega(n^{1/3} \ell_0^{2/3}(f_n))$ , then we show  $Q(f_n \square \wedge) = \Omega(\ell_1(f_n))$ . In both parts of the proof, we reduce an instance of  $f_{n'} \square \text{DISJ}_k^{\leq 1}$  to  $f_n \square \wedge$  for some appropriate function  $f_{n'}$  and  $k$ .

*Proof of Theorem 4.5.* Let  $c$  be the constant in Theorem 4.4,  $\beta \stackrel{\text{def}}{=} \min\{\sqrt[3]{2}, (\frac{c}{12e})^{2/3}\}$ , and  $\alpha \stackrel{\text{def}}{=} (\beta/2)^{3/2}$ .

Consider the case that  $\ell_0 \stackrel{\text{def}}{=} \ell_0(f_n) \leq \alpha n$ . Let  $n' \stackrel{\text{def}}{=} \beta n^{2/3} \ell_0^{1/3}$ , and  $f_{n'} \in \mathcal{F}_{n'}$  be such that  $f_{n'}(x) = f_n(x 0^{n-n'})$ ,  $\forall x \in \{0, 1\}^{n'}$ . By direct inspection,  $n' \leq n$ , thus  $f_{n'}$  is well-defined. Since

$$f_{n'}(1^{\ell_0-1} 0^{n'-\ell_0+1}) = f_n(1^{\ell_0-1} 0^{n-1\ell_0+1}) \neq f_n(1^{\ell_0} 0^{n-\ell_0}) = f_{n'}(1^{\ell_0} 0^{n'-\ell_0}),$$

and by direct inspection,  $\ell_0 \leq n'/2$ , we have  $\ell_0(f_{n'}) \geq \ell_0$ . By Theorem 4.4,

$$\widetilde{\deg}(f_{n'}) \geq c\sqrt{n'(\ell_0(f_{n'}) + \ell_1(f_{n'}))} \geq c\sqrt{n'\ell_0}.$$

Set  $k \stackrel{\text{def}}{=} \lceil \frac{6en'}{\deg(f_{n'})} \rceil$ . By Lemma 4.3,  $Q(f_{n'} \square \text{DISJ}_k^{\leq 1}) = \Omega(\widetilde{\deg}(f_{n'})) = \Omega(n^{1/3} \ell_0^{2/3})$ . Note that

$$n'k \leq \beta n^{2/3} \ell_0^{1/3} \cdot \frac{12e\sqrt{\beta}}{c} \left(\frac{n}{\ell_0}\right)^{1/3} = \beta^{3/2} \frac{12e}{c} n \leq n.$$

Therefore,  $\forall (x, y) \in \text{dom}(f_{n'} \square \text{DISJ}_k^{\leq 1})$ , we have  $(f_{n'} \square \text{DISJ}_k^{\leq 1})(x, y) = (f_n \square \wedge)(x 0^{n-n'k}, y 0^{n-n'k})$ . Thus  $Q(f_n \square \wedge) \geq Q(f_{n'} \square \text{DISJ}_k^{\leq 1}) = \Omega(n^{1/3} \ell_0^{2/3})$ .

Now consider the case that  $\alpha n < \ell_0 \leq n/2$ . Set  $k \stackrel{\text{def}}{=} \lceil \frac{6\sqrt{2}e}{c} \rceil$ , and  $n' \stackrel{\text{def}}{=} \min\{\frac{n-\ell_0+1}{2k-1}, \ell_0 - 1\}$ . Then  $n' = \Theta(n) = \Theta(\ell_0)$ . Define  $f_{n'} \in \mathcal{F}_{2n'}$  as follows:

$$f_{n'}(x) = f_n(x1^{\ell_0-1-n'}0^{n-2n'-(\ell_0-1-n')}), \quad \forall x \in \{0, 1\}^{2n'}.$$

By direct inspection,  $f_{n'}$  is well-defined. Then

$$f_{n'}(1^{n'}0^{n'}) = f_n(1^{\ell_0-1}0^{n-\ell_0+1}) \neq f_n(1^{\ell_0}0^{n-\ell_0}) = f_{n'}(1^{n'+1}0^{n'-1}).$$

Therefore,  $\ell_1(f_{n'}) = n'$ , and  $\widetilde{\deg}(f_{n'}) \geq \sqrt{2}cn'$ , by Theorem 4.4. By direct inspection,  $k \geq \frac{6e(2n')}{\deg(f_{n'})}$ , thus  $Q(f_{n'} \square \text{DISJ}_k^{\leq 1}) = \Omega(\widetilde{\deg}(f_{n'})) = \Omega(n')$ . Note that for all  $(x, y) \in \text{dom}(f_{n'} \square \text{DISJ}_k^{\leq 1})$ ,

$$(f_{n'} \square \text{DISJ}_k^{\leq 1})(x, y) = (f_n \square \wedge)(x1^{\ell_0-1-n'}0^{n-(\ell_0-1-n')-2kn'}, y1^{\ell_0-1-n'}0^{n-(\ell_0-1-n')-2kn'}).$$

By direct inspection, the number of 0's and 1's padded in the above equation is non-negative. Thus

$$Q(f_n \square \wedge) = \Omega(Q(f_{n'} \square \text{DISJ}_k^{\leq 1})) = \Omega(n') = \Omega(\ell_0) = \Omega(n^{1/3}\ell_0^{2/3}).$$

We use a similar reduction to prove  $Q(f_n \square \wedge) = \Omega(\ell_1)$ . Let  $k$  be the same as above. Set  $n' \stackrel{\text{def}}{=} \lfloor \frac{\ell_1}{2k-1} \rfloor$ , and define  $f_{n'} \in \mathcal{F}_{2n'}$  as follows

$$f_{n'}(x) = f_n(x1^{n-\ell_1-n'}0^{n-2n'-(n-\ell_1-n')}) \quad \forall x \in \{0, 1\}^{2n'}.$$

By direct inspection, the numbers of padded 0's and 1's are non-negative, thus  $f_{n'}$  is well-defined. Since

$$f_{n'}(1^{n'}0^{n'}) = f_n(1^{n-\ell_1}0^{\ell_1}) \neq f_n(1^{n-\ell_1+1}0^{n-\ell_1-1}) = f_{n'}(1^{n'+1}0^{n'-1}),$$

we have  $\ell_1(f_{n'}) = n'$ . Thus  $\widetilde{\deg}(f_{n'}) \geq \sqrt{2}cn'$  by Theorem 4.4, and  $Q(f_{n'} \square \text{DISJ}_k^{\leq 1}) = \Omega(\widetilde{\deg}(f_{n'})) = \Omega(\ell_1)$  by Lemma 4.3. For all  $(x, y) \in \text{dom}(f_{n'} \square \text{DISJ}_k^{\leq 1})$ ,

$$(f_{n'} \square \text{DISJ}_k^{\leq 1})(x, y) = (f_n \square \wedge)(x1^{n-\ell_1-n'}0^{n-2kn'-(n-\ell_1-n')}, y1^{n-\ell_1-n'}0^{n-2kn'-(n-\ell_1-n')}).$$

By direct inspection again, the numbers of the padded digits in the above are non-negative. Thus  $Q(f_n \square \wedge) \geq Q(f_{n'} \square \text{DISJ}_k^{\leq 1}) = \Omega(\ell_1)$ .  $\square$

Next, we establish a classical upper bound on the randomized complexity of symmetric predicates.

**Proposition 4.6.** *Let  $f_n \in \mathcal{F}_n$  be symmetric with  $\ell_0(f_n) = 0$ . Then*

$$R(f_n \square \wedge) = O(\ell_1 \log^2 \ell_1 \log \log \ell_1).$$

Theorem 1.7 follows from Theorem 4.5 and Proposition 4.6: if  $\ell_0(f_n) \geq 1$ ,  $Q(f_n \square \wedge) = \Omega(n^{1/3}) = \Omega(D^{1/3}(f_n \square \wedge) = \Omega(R^{1/3}(f_n \square \wedge))$ . Otherwise,  $Q(f_n \square \wedge) = \Omega(\ell_1(f)) = \Omega(R^{1/2}(f_n \square \wedge))$ . Similarly, Proposition 1.3 follows from Proposition 4.6 and Razborov's lower bound Equation 13.

To prove Proposition 4.6, we use the following result from Huang et al. [16]. Let  $n$  and  $d$  be integers with  $0 \leq d \leq n$ . The HAMMING DISTANCE PROBLEM  $\text{HAM}_{n,d}$  is defined as

$$\text{HAM}_{n,d}(x, y) = \begin{cases} 1 & |x \oplus y| \geq d, \\ 0 & \text{otherwise.} \end{cases}$$

**Theorem 4.7** (Huang et al. [16]). *There is randomized protocol for  $\text{HAM}_{n,d}$  that exchanges  $O(d \log d)$  bits and errs with probability  $\leq 1/3$ .*

*Proof of Proposition 4.6.* Without loss of generality, assume  $f_n(1^m 0^{n-m}) = 0$  for all  $m$ ,  $0 \leq m \leq n - \ell_1$ . The following randomized protocol computes  $f_n \square \wedge$  with  $O(\ell_1 \log^2 \ell_1 \log \log \ell_1)$  bits of communication. Fix an input  $(x, y)$ , and let  $z_A \stackrel{\text{def}}{=} n - |x|$  and  $z_B \stackrel{\text{def}}{=} n - |y|$ . Alice and Bob first check if  $z_A \geq \ell_1$  or  $z_B \geq \ell_1$ . If yes, they output 0 and terminate the protocol. Otherwise, Alice sends  $z_A$  to Bob using  $\lceil \log_2(\ell_1 - 1) \rceil$  bits, and they compute  $\delta \stackrel{\text{def}}{=} |x \oplus y|$ . Knowing  $z_A$  and  $\delta$ , Bob is able to compute  $f(|x \cap y|) = f((|x| + |y| - |x \oplus y|)/2)$ . Note that  $\Delta \stackrel{\text{def}}{=} 2(\ell_1 - 1) \geq \delta \geq 0$ . Thus Alice and Bob can perform a binary search to determine  $\delta$  with  $\log_2(\Delta + 1)$  sub-protocols for the HAMMING DISTANCE PROBLEM. For each candidate value  $d$  of  $\delta$ , they repeat the randomized protocol in Theorem 4.7 for  $\text{HAM}_{n,d}$  for  $\Theta(\log \log \Delta)$  times so that the error probability is  $\leq \frac{1}{3(\log_2 \Delta + 1)}$ . Thus the total number of bits exchanged is  $O(\Delta \log^2 \Delta \log \log \Delta) = O(\ell_1 \log^2 \ell_1 \log \log \ell_1)$ , and the error probability of the complete protocol is  $\leq 1/3$ .  $\square$

*Remark 4.8.* While both Razborov’s proof and the above use the spectrum decompositions of the matrix  $J_{k,p,s}$ , we emphasize their difference: we only need to analyze  $\|\frac{\mu_0 - \mu_1}{2}\|$ , which corresponds to  $s = 0, 1$ . In contrast, Razborov’s proof needs much more details of the spectrum decompositions, in particular, it needs to consider  $s = 0, 1, \dots, \Theta(n)$ .

Theorem 1.7 implies  $Q(\text{DISJ}_n) = \Omega(n^{1/3})$ . Note that our estimate (Equation 12) gives  $\rho(\text{DISJ}_k) = O(1/k)$ . Thus by Proposition 3.4, this only gives a very weak lower bound  $Q(\text{DISJ}_n) = \Omega(\log n)$ . Surprisingly, this weak bound can be amplified to  $\Omega(n^{1/3})$  through the dual formulation of the approximate degree (Lemma 3.1). Finding more examples of such “hardness amplification” would be very interesting.

## 5 Open problems and discussions

While the block-composed functions we focus on are restricted to have identical  $g_k$  in each block, and  $g_k$  has balanced input size on Alice and Bob’s side, our technique can be extended straightforwardly to deal with non-identical, and general building block functions. Pushing this approach to its limit in resolving the Log-Equivalence Conjecture is an interesting direction.

A specific problem is to minimize the technical assumption on the block-size in the Main Lemma — for some  $g_k$ , this can be accomplished by using the result of Sherstov [33], which we will describe below in more details. Another specific problem is to prove the Log-Equivalence Conjecture for  $f_n \square \wedge$ , for an arbitrary  $f_n$ .

In an independent work, Sherstov [33] also derived Lemma 3.1, and used it to prove strong quantum lower bounds on what he called “pattern matrices”. In our notation, he considered functions  $f_n \square g_k^0$ , where  $f_n \in \mathcal{F}_n$  and  $g_k^0 : \{0, 1\}^k \times ([k] \times \{0, 1\}) \rightarrow \{0, 1\}$  is fixed with  $g_k^0(x, (i, b)) \stackrel{\text{def}}{=} x_i + b$ . His main result is,  $Q(f_n \square g_k^0) = \Omega(\deg(f_n))$  for any  $f_n$ . The proof also starts with the dual characterization of  $\deg(f_n)$ , constructs  $q$  via Lemma 3.1, then constructs a witness matrix  $h$  (or  $K$  in [33]) for the high trace norm of any matrix approximating  $f_n \square g_k^0$ . His construction of  $h$  can be expressed in the same equation (Eqn. 3) as ours with carefully chosen  $\mu_0$  and  $\mu_1$  for  $g_k^0$ .

The main technical difference takes place after Eqn. (5). With the fixed  $g_k^0$ , the constructed  $h$  has the nice property that the left and right eigenvectors of  $(\mu_0 + \mu_1)^{\otimes \bar{w}} \otimes (\mu_0 - \mu_1)^{\otimes w}$  are in

orthogonal subspaces, due to the fact that

$$(\mu_0 + \mu_1)^T(\mu_0 - \mu_1) = 0, \quad \text{and,} \quad (\mu_0 - \mu_1)^T(\mu_0 + \mu_1) = 0. \quad (14)$$

Thus, he was able to avoid the use of the triangle inequality in Eqn. (6) and replace the summation by the maximum. This sharper bound moderates the requirement on  $k$ , and results in an alternative proof for Razborov’s lower bound with the same asymptotic parameters and without using Hahn polynomials at all. In particular, he proved that  $Q(f_n \square \text{DISJ}_k) = \Omega(\widetilde{\deg}(f_n))$  for any  $f_n$  and any  $k \geq 4$ . This is a significantly stronger result than our requirement that  $k \geq \frac{6en}{\deg(f_n)}$  (Lemma 4.3) when  $\widetilde{\deg}(f_n)$  is much smaller than  $n$ . On the other hand, for a general  $g_k$ , the best bound on  $Q(f_n \square g_k)$  provable through this method (i.e., using pairs of  $\mu_0$  and  $\mu_1$  satisfying the orthogonality condition (14)), is not necessarily stronger than that in Main Lemma. This is because the orthogonality condition restricts the choice of  $\mu_0$  and  $\mu_1$  to smaller domains.

## 6 Acknowledgments

We thank Jianxin Chen, Sasha Razborov, Sasha Sherstov, and Zhiqiang Zhang for useful discussions. We also thank Sasha Sherstov for sending us his manuscript [33].

## References

- [1] Aaronson and Ambainis. Quantum search of spatial regions (extended abstract). In *Proceedings of the 44th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 200–209, 2003.
- [2] A. Ambainis, L. Schulman, A. Ta-Shma, U. Vazirani, and A. Wigderson. Quantum communication complexity of sampling. In *Proceedings of 39th Annual Symposium on the Foundations of Computer Science (FOCS)*, pages 342–351, 1998.
- [3] Z. Bar-Yossef, T. S. Jayram, and I. Kerenidis. Exponential separation of quantum and classical one-way communication complexity. In *Proceedings of the thirty-sixth annual ACM Symposium on Theory of Computing (STOC)*, pages 128–137, 2004.
- [4] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. In *Proceedings of the 39th Annual Symposium on Foundations of Computer Science*, pages 352–361, 1998.
- [5] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. *Journal of the ACM*, 48(4):778–797, 2001.
- [6] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf. Quantum fingerprinting. *Phys. Rev. Lett.*, 87(16):167902, 2001.
- [7] H. Buhrman, R. Cleve, and A. Wigderson. Quantum vs. classical communication and computation. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 63–68, 1998.

- [8] H. Buhrman and R. de Wolf. Communication complexity lower bounds by polynomials. In F. M. Titsworth, editor, *Proceedings of the 16th Annual Conference on Computational Complexity (CCC)*, pages 120–130, 2000.
- [9] H. Buhrman and R. de Wolf. Complexity measures and decision tree complexity: a survey. *Theoretical Computer Science*, 288(1):21–43, 2002.
- [10] R. Cleve. The query complexity of order-finding. *Information and Computation*, 192(2):162–171, 2004.
- [11] R. Cleve, W. van Dam, M. Nielsen, and A. Tapp. Quantum Entanglement and the Communication Complexity of the Inner Product Function. *Lecture Notes in Computer Science*, 1509:61–74, 1999.
- [12] P. Delsarte. Hahn polynomials, discrete harmonics and t-designs. *SIAM Journal on Applied Mathematics*, 34(1):157–166, 1978.
- [13] D. Gavinsky, J. Kempe, and R. de Wolf. Exponential separation of quantum and classical one-way communication complexity for a Boolean function. quant-ph/0607174.
- [14] D. Gavinsky, J. Kempe, O. Regev, and R. de Wolf. Bounded-error quantum state identification and exponential separations in communication complexity. In *Proceedings of the 38th Symposium on Theory of Computing (STOC)*, pages 594–603, 2006.
- [15] P. Høyer and R. de Wolf. Improved quantum communication complexity bounds for disjointness and equality. *Lecture Notes in Computer Science*, 2285:299–310, 2002.
- [16] W. Huang, Y. Shi, S. Zhang, and Y. Zhu. The communication complexity of the Hamming Distance Problem. *Information Processing Letters*, 99(4):149–153, 2006.
- [17] I. Kerenidis and R. Raz. The one-way communication complexity of the boolean hidden matching problem. quant-ph/0607173.
- [18] H. Klauck. Lower bounds for quantum communication complexity. In *Proceedings of the 42nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 288–297, 2001.
- [19] H. Klauck. Rectangle size bounds and threshold covers in communication complexity. In *Proceedings of the 18th IEEE Conference on Computation Complexity*, pages 118–134, 2003.
- [20] D. Knuth. Combinatorial matrices. Manuscripts available at <http://www-cs-faculty.stanford.edu/~knuth/preprints.html>, 1991.
- [21] I. Kremer. *Quantum Communication*. Master’s Thesis, Computer Science Department, The Hebrew University of Jerusalem, 1995.
- [22] I. Kremer, N. Nisan, and D. Ron. On randomized one-round communication complexity. In *Proceedings of the twenty-seventh annual ACM symposium on Theory of computing (STOC)*, pages 596–605, 1995.
- [23] E. Kushilevitz and N. Nisan. *Communication Complexity*. Cambridge University Press, Cambridge, 1997.



- [24] T. Lee, A. Shraibman, and R. Špalek. Disjointness is hard in the multi-party number-on-the-forehead model. To appear in *Proceedings of the 23rd Annual Conference on Computational Complexity (CCC)*, 2008.
- [25] N. Linial and A. Shraibman. Lower bounds in communication complexity based on factorization norms. *Proceedings of the Thirty-Ninth Annual ACM Symposium on the Theory of Computation (STOC)*, pages 669–708, 2007.
- [26] N. Linial and A. Shraibman. Learning complexity vs. communication complexity. To appear in *Proceedings of the 23rd Annual Conference on Computational Complexity (CCC)*, 2008.
- [27] N. Nisan and M. Szegedy. On the degree of Boolean functions as real polynomials. In *Proceedings of the Twenty-Fourth Annual ACM Symposium on the Theory of Computing (STOC)*, pages 462–467, 1992.
- [28] R. Paturi. On the degree of polynomials that approximate symmetric Boolean functions (preliminary version). In *Proceedings of the Twenty-Fourth Annual ACM Symposium on the Theory of Computing (STOC)*, pages 468–474, 1992.
- [29] R. Raz. Exponential separation of quantum and classical communication complexity. *Proceedings of the 31st ACM Symposium on Theory of Computing (STOC)*, pages 358–367, 1999.
- [30] A. A. Razborov. Personal communication, 2002.
- [31] A. A. Razborov. Quantum communication complexity of symmetric predicates (Russian). *Izvestiya: Mathematics*, 67(1):145–159, 2002. English translation available at [http://genesis.mi.ras.ru/razborov/qcc\\_eng.ps](http://genesis.mi.ras.ru/razborov/qcc_eng.ps).
- [32] A. A. Sherstov. Separating  $AC^0$  from depth-2 majority circuits. In *Proceedings of the 39th Symposium on Theory of Computing (STOC)*, pages 294–301, 2007.
- [33] A. A. Sherstov. The pattern matrix method for lower bounds on quantum communication. Technical report TR-07-46, The Univ. of Texas at Austin, Dept. of Comp. Sci., 6 Sept. 2007.
- [34] D. R. Simon. On the power of quantum computation. *SIAM Journal on Computing*, 26(5):1474–1483, 1997.
- [35] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997.
- [36] M. Szegedy. Personal communication, 2004.
- [37] A. C.-C. Yao. Some complexity questions related to distributive computing. In *Eleventh Annual ACM Symposium on Theory of Computing (STOC)*, pages 209–213, 1979.
- [38] A. C.-C. Yao. Quantum circuit complexity. In *Proceedings of the 34th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 352–361, 1993.